NASA Taxonomy 2.0 Progress Report
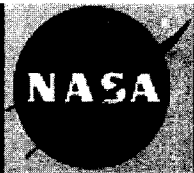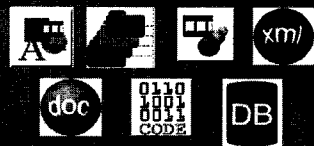
Jayne Dutra, Jet Propulsion Laboratory
Joseph Busch, Taxonomy Strategies
NASA CIO Office
February 10, 2004
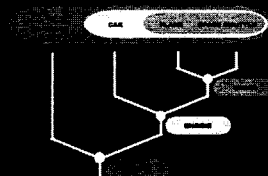
Purpose of the NASA Taxonomy

Discover

FIRSTGOV

NASA NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

Classify

Create

- Site Maps
- Search Engine
- NASA Portals
- Content Integration Networks

- Logical & Intuitive Filters
- Taxonomy

- Content
- Assets

Finding the right information at the right time to solve the problem at hand

JPL Knowledge Management/
Taxonomy Strategies
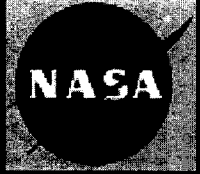
Project Benefits: Enable Knowledge Discovery

JPL

- Make it easy for various audiences to find relevant information from NASA programs quickly
  - Provide easy access for NASA resources found on the Web
  - Share knowledge by enabling users to easily find links to databases and tools
  - Provide search results targeted to user interests
  - Enable the ability to move content through the enterprise to where it is needed most

- Comply with E-Government Act of 2002
- Be a leading participant in federal XML projects

3
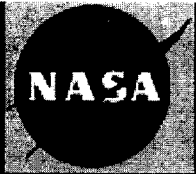
- # Design process that:
  - Incorporates existing federal and industry terminology standards like NASA AFS, NASA CMS, FEA BRM, NAICS, and IEEE LOM
  - Provides a product for the NASA XML namespace registry
  - Complies with metadata standards like Z39.19, ISO 2709, and Dublin Core

- # Practices increase interoperability and extensibility

Project Progress: NASA Taxonomy 2.0 Revision Completed

JPL

Follow-on Work
> Integrate with applications.

Phase 4
> Dublin Core mapping
> XML schema development

Phase 3
> Test & validate Taxonomy

Phase 2
> Build community of practice
> Agree on comprehensive branches & taxonomy detail

Phase 1
> Audit Content practices
> Identify & survey stakeholder

PROJECT 1

PROJECT 2

JPL Knowledge Management/ Taxonomy Strategies

5

- Understand current strategies and practices for creating, collecting, and organizing information across NASA
- Observe how information is used and organized, the audiences for this information, and the information needs of these audiences
- Elicit goals, hopes, and concerns for an information architecture solution
- Start building a community of interest

6
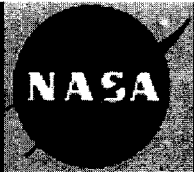
Key Findings

- Most (70%) NASA content already has some tagging or is categorized.
- Most (70%) owners add tag content with metadata.
- Almost half (45%) use a standard metadata schema, but many different standards are used.
- Most (60%) use a controlled vocabulary, but lots of different controlled vocabularies are used.

➢ Different NASA constituencies care deeply about what schemas are specified, and what vocabularies are used because ...

*JPL Knowledge Management/*
Taxonomy Strategies

Admin

- Better understand the program in total, and obtain scheduling information, project status and best practices.
- Access procurement rules and examples, and procurement action synopses.

Sci Tech

- Engineering specifications.
- Scholarly research, competitive intelligence, and general aerospace research.
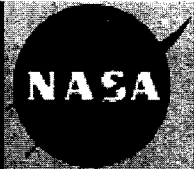
Public

- Develop educational products, support current products, learn, etc.
- Topic research and fact finding, topic background research, and downloading curriculum support materials.
- In the classroom as stand alone items, hands-on learning opportunities, class projects, to expand on a student's learning potential.
- Find NASA contact information on services, information about student opportunities, information about career opportunities, and latest educational news.
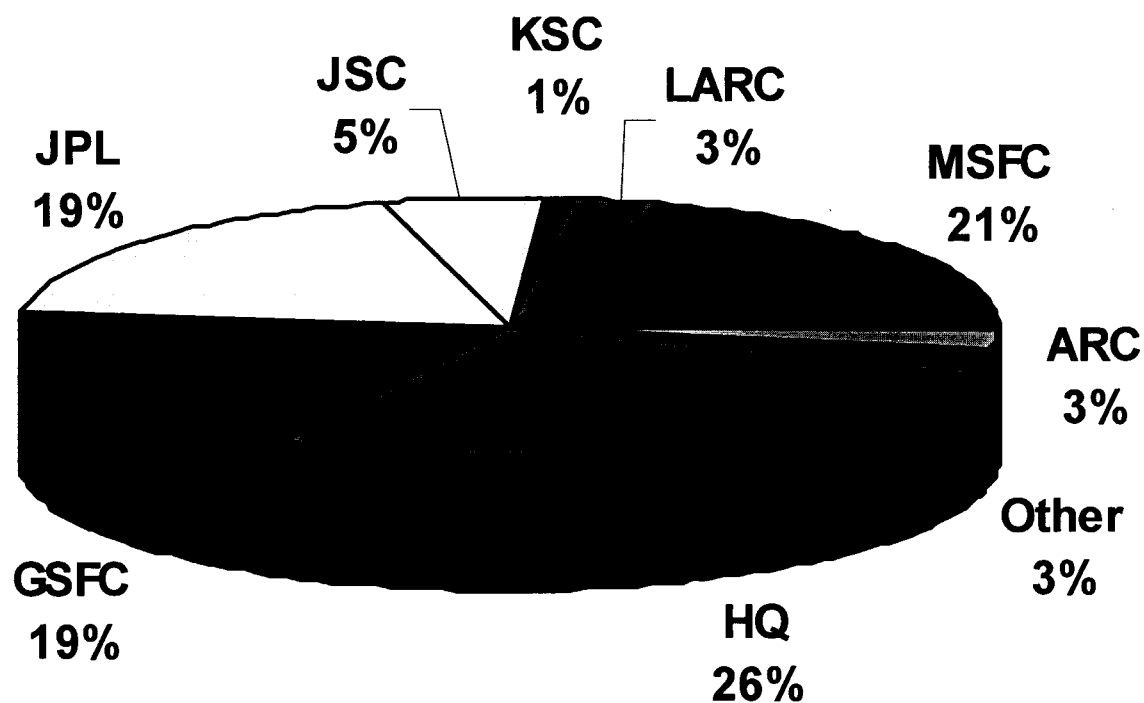
JPL
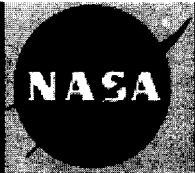
Test and Validation Phase

- Qualitative validation
  - Confirm stakeholders and communities
  - Focus on Projects, Engineering & Science
- Quantitative validation
  - Select and build test collection
  - Stratify automated categorizer
- Extend taxonomy value space as needed
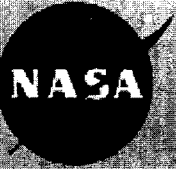- Review results with stakeholders and report to CIO Council

9

JPL

Pie chart "by location":
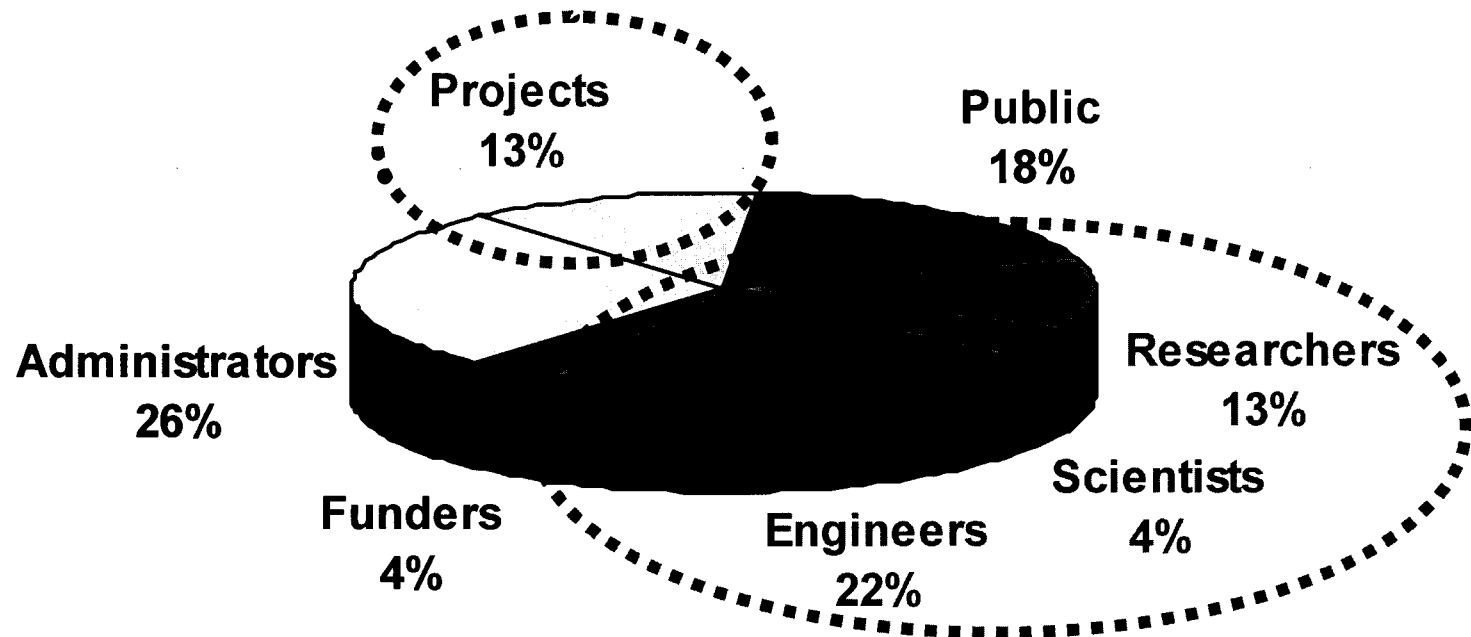- JSC 5%
- KSC 1%
- LARC 3%
- JPL 19%
- MSFC 21%
- ARC 3%
- Other 3%
- HQ 26%
- GSFC 19%

... by location

JPL Knowledge Management/
Taxonomy Strategies

Focused on Projects, Engineering & Science

JPL

Projects 13%

Public 18%

Administrators 26%

Researchers 13%

Scientists 4%

Funders 4%

Engineers 22%

... by primary audience served.

52%-Projects, Engineering & Science

JPL Knowledge Management/
Taxonomy Strategies

JPL

Collection	Source URL	No of Docs
Lessons Learned Database	http://llis.nasa.gov	1,370
NTRS (NASA Technical report Server)	http://ntrs.nasa.gov	213,900
SIRTF (Space Infrared Telescope Facility) Project Library	http://sirtifweb.jpl.nasa.gov	4,054
James Webb Space Telescope (JWST) Project Documents	http://ngst.gsfc.nasa.gov/doclist/bytitle.html	634

12
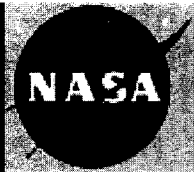
Built Demonstration of Taxonomy Value in Search and Navigation

JPL

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

Search | rover | [Go]

219958 items

by Organization	by Subject	by Missions and Projects	by Date
NASA Affiliated Institutions 1378	Aeronautics 26532	Aerospace Technology 60	1972 8392
NASA Centers 76545	Astronautics 31758	Biological and Physical Research 68	1973 8512
NASA Contractors 10108	Chemistry and Materials 17086	Data 140	1974 7828
NASA Enterprises 815	Engineering 39631	Earth Sciences 1497	1975 7704
NASA Headquarters 4042	Geosciences 30770	Human Exploration and Development... 10680	1992 8131
Other NASA Partners 999	Mathematical and Computer Sciences 13286	Planetary Missions 4819	1993 8519
	Space Sciences 22685	Space Sciences 9467	1994 7712
	4 more		*74 more*

by Competencies	by Information Type	by Collection
Business 386	Catalogs and Databases 32	LessonsLearned 1370
Engineering 393	Designs and Specifications 62	NTRS 213900
Mission 555	Plans and Agendas 158	SIRTF 4054
Scientific 410	Results and Analyses 260	Webb 634
Technical 218	Reviews and Lessons Learned 1819	
	Status Reports 119	
	Technical Reports 229	
	6 more	

...that provides common access framework across test collections
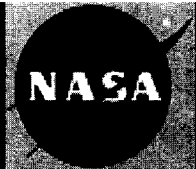
by Organization
- NASA Affiliated Institutions 6
- NASA Centers 304
- NASA Contractors 18
- NASA Enterprises 3
- NASA Headquarters 7
- Other NASA Partners 14

by Subject
- Aeronautics 13
- Astronautics 109
- Engineering 106
- Geosciences 26
- Life Sciences 31
- Mathematical and Computer Sciences 65
- Space Sciences 208
- *4 more*

by Missions and Projects
- Data 1
- Earth Sciences 3
- Human Exploration and Development... 130
- Planetary Missions 125
- Space Sciences 7

by Date
- 1989 37
- 1991 40
- 1992 44
- 1993 44
- 1999 61
- 2000 60
- 2001 40
- *29 more*

by Competencies
- Business 4
- Engineering 4
- Mission 3
- Scientific 3

by Information Type
- Reviews and Lessons Learned 5
- Status Reports 1
- Technical Reports 3

by Collection
- LessonsLearned 4
- NTRS 591

14

5 items matching
Text contains rover

Information Type:
Reviews and Lessons
 Learned

by Organization
- NASA Centers 5
- NASA Enterprises 3

by Subject
- Aeronautics 4
- Astronautics 4
- Chemistry and Materials 1
- Engineering 1
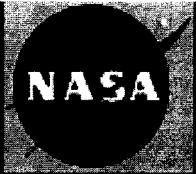- Mathematical and Computer Sciences 1

by Date
- 1997 1
- 2002 1
- 2003 3

by Competencies
- Business 4
- Engineering 4
- Mission 3
- Scientific 3

by Collection
- Lessons Learned 4
- NTRS 1

15

JPL

- http://spongebob.siderean.com:7880/test /test2query3.jsp
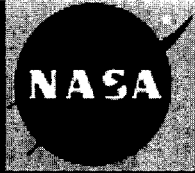
  Logon: NASA
  Password: facets

  - Hosted by Seamark www.seamark.com with Siderean software

16

Taxonomy Basics

What is the NASA Taxonomy?

- The classification scheme is meant to encompass all of NASA web content (NASA web space) including internal as well as external material. It is a means for tagging content so it can be used and reused in different contexts.

How to Use the NASA Taxonomy

- This is a generic taxonomy from which specializations can be derived for specific purposes.
    - **Not all facets need to be used in each instance**
    - **A facet is repeatable**
   - **The taxonomy is modular and dynamic**
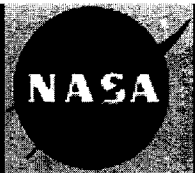
Extend Taxonomy Value Space as Needed

NASA Taxonomy Facets (Top Level)

- Audiences
- Business Purpose (*formerly Functions*)
- Competencies (*formerly Disciplines*)
- Content Types (*formerly Information*)
- Industries
- Instruments (*new*)
- Locations
- Missions and Projects
- Organizations
- Subject Categories (*new*)

- Dates (*formerly Chronology*)
- Collections

http://eis.jpl.nasa.gov/~jedutra/NASA_Taxonomy_04/index.htm

JPL Knowledge Management/
Taxonomy Strategies
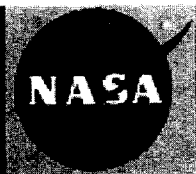
- **Content Types**
  - . Announcements
  - . . Press Kits
  - . . Press Releases
  - . Articles, Notes, and Papers
  - . Calendars and Schedules
  - . . Agendas
  - . Case Studies
  - . Catalogs and Databases
  - . Correspondence
  - . . e-Mails
  - . . Memos
  - . Databases
  - . . Bibliographic Databases
  - . . Image Databases
  - . Designs and Specifications
  - . . Configuration Controls
  - . . Notebooks
  - . . Quality Control
  - . . Requirements
  - . Drawings
  - . Educational Materials
  - . . Activity Guides
  - . . Educational Toys
  - . . Educator's Guides

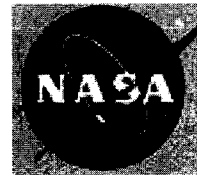**Configuration Controls**

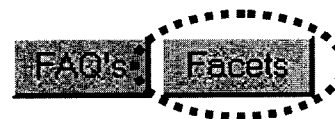**Broader Terms:**
Designs and Specifications

Scope Note:
Records of changes to documentation or hardware, including engineering change requests and waivers.

Term Number:
52

19

■ ■ ■

JPL

NASA Taxonomy

FAQ's   Facets

Search FirstGov   Search Google   A-Z Index   Facet Hierarchy   Term Index

Context-sensitive comments email

20

JPL

- http://eis.jpl.nasa.gov/~jedutra/NASA_Ta xonomy_04/index.htm

---

- Taxonomy Website built using MultiTes software www.multites.com
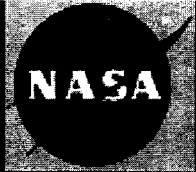
JPL
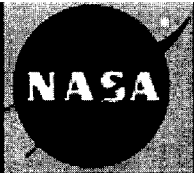
Dublin Core Mapping and XML Schema Development

- Complete Dublin Core mapping
- Create any necessary NASA specific tags
- Develop XML schema from metadata
- Review results with stakeholders and report to CIO Council
- Register schemas in NASA XML Registry
- Educate and train publishing communities

JPL

- NASA Taxonomy provides controlled vocabularies used to populate elements of more complex metadata schema such as the Dublin Core (www.dublincore.org)

- The taxonomy facets map to these schema elements.

23

NASA Taxonomy – Dublin Core Map (Draft)

| Dublin Core Elements	Definition	NASA Taxonomy Mapping
Creator	Content maker.	dc:creator dc:creator.employee dc:creator.organization
Subject	Content topic.	dc.subject.organization dc:subject.missionsProjects dc:subject.disciplines
Publisher	Publisher of this manifestation.	dc:publisher.organization
Contributor	Content contributor.	dc:contributor dc:contributor.employee dc:contributor.organization
Type	Genre.	dc:type.information
Coverage	Space, period, date, jurisdiction, etc.	dc:coverage.locations dc:coverage.chronology
Audience	Content audience.	dcTERM:audience
Non DC	NASA missions and projects.	nasa:missionsProjects
Non DC	Business functions.	nasa:functions
Non DC	Technical specialties.	nasa:disciplines
Non DC	Standard industry categories.	naics:industries

... at the NASA Level

- Metadata specification for all NASA content publishers
- XML schema registered in accordance with agreed standards (to enable appropriate use and reuse)
- Enhancement of Agency Web publishing processes
- Integration with NASA public portal content management system for:
  - Reduced publishing cycles
  - Better quality of Web materials – coordinated themes
- Integration with NASA Search Engine, Web Site Registration System
- Application in many technical areas, including engineering and science disciplines (STEP and science data dictionaries)
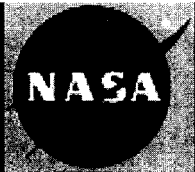
Project Outcomes: NASA Taxonomy Benefits

... at the Federal Level

- NASA taxonomy development in accordance with e-Gov Act of 2002
- Integration with FEA at the BRM & DRM level
- Increased interoperability with other federal agencies through common data models and standards
- Better interoperability with industry partners for increased speed of mission development
- Enhanced results in First Gov search engine
- Readiness to actively participate in e-Gov initiatives
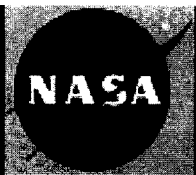
- Taxonomy stewardship
  - Maintenance, education and training
  - Facilitate standard adoption process
- Apply in public and internal portals
  - DM, PM, CM, and DAM tagging
  - Search integration
    - Verity K2
    - Faceted search and navigation
  - content integration networks for real time delivery
- Implement in website registration system
- Participate in NASA Enterprise Architecture Group standards development
  - Agency UDDI, RSS Syndication

27

- CIO presentation
- Define review and approval process
- Facilitate review and approval process
- Formalize taxonomy as XML schema
- Delegate taxonomy ownership
  - Within Code AO
- Plan follow-on work

JPL Knowledge Management/
Taxonomy Strategies

JPL

- NASA Taxonomy Report and Recommendations, Dutra, Busch and Daniels, 2004
  - url

- NASA Taxonomy White Paper, Dutra and Busch, 2003
  - https://pub-lib.jpl.nasa.gov/pub-lib/dscgi/ds.py/Get/File-22/NASA_Taxonomy_Overview-0203.ppt

- Implementing the NASA Taxonomy Through Service Oriented Architectures, Dutra and Xiao, 2004
 - url

Taxonomy Tagging

Examples

Taxonomy Example

http://wufs.wustl.edu/missions/odyssey/
#Odyssey%20Data%20Sets
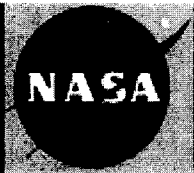
Attribute	Values
Content Types	Data Files; Web Sites
Audiences	Researchers; Scientists
Organizations	Jet Propulsion Laboratory
Missions and Projects	Mars Odyssey
Industries	N/A
Locations	Mars
Business Purpose	Scientific and Technical Information
Competencies	Planetary and Lunar Science
Dates	2002-present

31
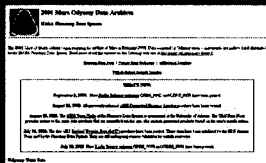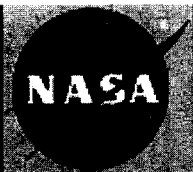
Clementine – DSPSE

Taxonomy Example

http://www.cmf.nrl.navy.mil/clementine/
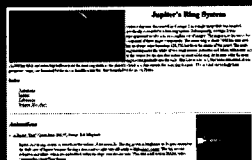
Attribute	Values
Content Types	Web Sites; Data Files; Images
Audiences	Researchers; Scientists; Educators; Students
Organizations	Naval Research Laboratory
Missions and Projects	Clementine
Industries	N/A
Locations	The Moon
Business Purpose	Scientific and Technical Information
Competencies	Planetary and Lunar Science
Dates	1994

JPL Knowledge Management/
Taxonomy Strategies
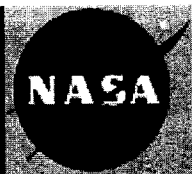
Jupiter's Ring System

Taxonomy Example

http://ringmaster.arc.nasa.gov/jupiter/
jupiter.html#index

Attribute	Values
Content Types	Web Sites; Animations; Images; Reference Sources
Audiences	Educators; Students
Organizations	Ames Research Center
Missions and Projects	Voyager; Galileo; Cassini; Hubble Space Telescope
Industries	N/A
Locations	Jupiter
Business Purpose	Scientific and Technical Information
Competencies	Planetary and Lunar Science
Dates	1979-1999

*JPL Knowledge Management/
Taxonomy Strategies*